

Спектр структурных вариаций генома у больных с ишемической болезнью сердца

Слепцов А.А.¹, Назаренко М.С.^{1,2,3}, Скрыбин Н.А.¹, Казанцев А.Н.², Барбараш О.Л.², Пузырев В.П.^{1,3}

¹ Научно-исследовательский институт медицинской генетики, Томский национальный исследовательский медицинский центр Российской академии наук, г. Томск, Россия

² Федеральное государственное бюджетное научное учреждение «Научно-исследовательский институт комплексных проблем сердечно-сосудистых заболеваний», г. Кемерово, Россия

³ Федеральное государственное бюджетное образовательное учреждение высшего образования «Сибирский государственный медицинский университет» Министерства здравоохранения Российской Федерации, г. Томск, Россия
e-mail: alexei.sleptcov@medgenetics.ru

Предполагается, что вариации числа копий участков ДНК (copy number variation, CNV) могут вносить вклад в формирование генетической структуры многофакторных заболеваний, тем самым объясняя определенную долю их наследуемости. Вместе с тем, поиск CNV у больных с ишемической болезнью сердца (ИБС) с использованием матричной сравнительной геномной гибридизации (array comparative genome hybridization, aCGH) ранее не проводился. Цель исследования заключалась в оценке спектра и характеристике CNV у больных с ИБС с использованием технологии aCGH. Скрининг CNV выполнен с применением высокопроизводительных микрочипов SurePrint G3 Human CGH+SNP 2x400 K (Agilent Technologies). Тестируемые образцы ДНК получены из лейкоцитов периферической крови мужчин с ИБС (n = 10). В качестве референсной использовалась ДНК мужчины европейского происхождения (Agilent Euro Male, Agilent Technologies). Всего в лейкоцитах идентифицировано 90 CNV, среди них – 72 (80%) содержат гены, белковые продукты которых участвуют в иммунновоспалительном ответе, обеспечивают функционирование обонятельных рецепторов, а также ферментов метаболизма. Гены, картированные в области CNV в хромосомных субсегментах 1p22.2 (*GBP3*), 1p21.1 (*AMY2B*) и 22q11.23 (*GSTT1*, *LOC391322*), ранее были связаны с атеросклерозом и его факторами риска.

Ключевые слова: вариации числа копий участков ДНК, ишемическая болезнь сердца.

Авторы декларируют отсутствие конфликта интересов.

Исследование было проведено при поддержке гранта Российского научного фонда №14-15-00305.

Spectrum of structural variations in patients with coronary heart disease

Sleptsov A.A.¹, Nazarenko M.S.^{1,2,3}, Skryabin N.A.¹, Kazantsev A.N.², Barbarash O.L.², Puzyrev V.P.^{1,3}

¹ Research Institute of Medical Genetics, Tomsk National Research Medical Center of the Russian Academy of Sciences, Tomsk, Russian Federation

² Federal State Budgetary Scientific Institution Research Institute for Complex Issues of Cardiovascular Diseases, Kemerovo, Russian Federation

³ Siberian State Medical University, Tomsk, Russian Federation
e-mail: alexei.sleptcov@medgenetics.ru

It is expected that copy number variation (CNV) potentially contribute to the formation of a genetic structure of complex diseases thus they would explain a part of the missing heritability. However, a search for CNV in patients with coronary heart disease (CHD) using high-resolution technique of array comparative genome hybridization (aCGH) has not previously been conducted. The goal of our study was to evaluate CNV spectrum and characteristics in patients with CHD using aCGH. The CNV screening was performed using high-resolution microarrays SurePrint G3 Human CGH + SNP 2x400 K (Agilent Technologies). The DNA samples were obtained from peripheral blood leukocytes of men with CHD (n = 10). Agilent Euro male DNA was used as a reference. We identified 90 CNV, among them 72 (80%) contained genes encode proteins related to the immune and inflammatory response, activity of olfactory receptors and metabolic enzymes. Genes mapped in the CNV region in the 1p22.2 (*GBP3*), 1p21.1 (*AMY2B*) and 22q11.23 (*GSTT1*, *LOC391322*), were previously reported as associated with atherosclerosis and its risk factors.

Key words: copy number variation, coronary heart disease.

Введение

Определение вклада генетической компоненты в риск развития ишемической болезни сердца (ИБС) является одной из приоритетных задач предиктивной и прецизионной медицины. Проведение широкогеномных ассоциативных исследований (Genome-Wide Association Study, GWAS) позволило идентифицировать бо-

лее 150 локусов, ассоциированных с ИБС, 46 из которых подтвержались в нескольких работах [1]. В то же время, анализируя только один тип генетической изменчивости, в частности однонуклеотидный полиморфизм (single nucleotide polymorphism, SNP), можно объяснить лишь малую долю наследуемости патологии [2]. Предполагается, что определенный вклад в риск развития и

течение ИБС вносят вариации числа копий участков ДНК (copy number variation, CNV).

Данный тип структурных вариаций характеризуется вариабельностью участков ДНК по количеству их копий в геноме [3]. Такие вариации могут быть описаны с позиции референсного генома как CNV-амплификации или увеличение числа копий участков ДНК и CNV-делеции — уменьшение числа их копий. В зависимости от своего размера и локализации, данные варианты могут изменять структуру нуклеотидной последовательности регуляторных элементов и генов, влияя на функциональную активность генов, изменяя функцию белков. Предполагается, что редкие и крупные CNV обладают большим эффектом на фенотип, тогда как малые и частые вносят меньший вклад в его развитие [4].

Существует небольшое количество работ, связывающих CNV с ИБС или её факторами риска. Факторами риска развития ИБС является изменение числа копий нуклеотидной последовательности в гене *LPA*, что детерминирует уровень его белкового продукта в сыворотке крови индивидов, а CNV в гене *LDLR* приводит к развитию семейной гиперхолестеринемии [3]. У мужчин из когорты Northwick Park Heart Study II (Великобритания) идентифицированы CNV в генах *TLR4*, *SREBP1* и *IL-6ST*, которые показали ассоциацию с концентрацией аполипопротеинов AI и B [6]. Семь CNV, картированных в локусах 1p21.3, 1q31.2 (*CDC73*), 1q42.2 (*DISC1*), 3p21.31 (*CDCP1*), 10q11.21 (*RET*) 12p12.3 (*PIK3C2G*) и 16q23.3 (*CDH13*), связаны с ИБС в сочетании с гиперлипидемией у населения о. Тайвань [5]. В то же время, в работе другого исследовательского коллектива не установлено вклада CNV в риск развития атеросклероза коронарных артерий и инфаркта миокарда с ранней манифестацией [7]. Различные результаты данных работ могут быть связаны с популяционными особенностями выборок, а также различиями в отношении изучаемых клинических форм данного заболевания. Кроме того, в данных работах для анализа CNV использовались микрочипы к однонуклеотидному полиморфизму.

В настоящее время широкое распространение получила матричная сравнительная геномная гибридизация (array comparative genome hybridization, aCGH) с использованием которой можно в ходе одного эксперимента осуществить полногеномный скрининг CNV. Микрочипы SurePrint G3 Human CGH+SNP 2x400 K (Agilent Technologies) позволяют идентифицировать относительно небольшие по размеру вариации числа копий участков ДНК со средним разрешением около 7,2 т.п.н. Не исключено, что использование технологии aCGH выявит структурные вариации в тех регионах генома, которые ранее не были проанализированы с использованием микрочипов к однонуклеотидным вариантам.

Цель настоящего исследования заключалась в оценке спектра и характеристике CNV у больных с ИБС с использованием технологии aCGH.

Материалы и методы

В исследуемую группу вошли 10 мужчин с ИБС, славянского происхождения. Средний возраст индивидов составил $60,7 \pm 7,9$ года. Все мужчины имели выраженное атеросклеротическое поражение коронарных артерий, артериальную гипертензию, гиперхолестеринемия, сахарный диабет 2 типа и инфаркт миокарда в анамнезе. Формирование выборок и клиническая характеристика больных проводили на базе ФГБНУ «Научно-исследовательский институт комплексных проблем сердечно-сосудистых заболеваний» (г. Кемерово). Все индивиды подписали добровольное информированное согласие на участие в исследовании. Проведение данного исследования одобрено этическим комитетом НИИ медицинской генетики Томского НИМЦ.

Выделение ДНК из лейкоцитов периферической крови проводилось с применением наборов QIAamp DNA Micro Kit (Qiagen), согласно протоколу производителя. Оценка качества и концентрации выделенных образцов ДНК проведена на спектрофотометре NanoDrop (Thermo Scientific) и путём электрофореза в 1% агарозном геле.

Скрининг CNV в образцах ДНК 10 пациентов выполнен методом aCGH на микрочипах SurePrint G3 Human CGH+SNP 2x400K (G4842A, Agilent Technologies), согласно протоколу производителя. В качестве референсной ДНК использовалась «ДНК мужчины европейского происхождения» (Agilent Euro Male, # 5190, Agilent Technologies). Сканирование микрочипов проводилось на приборе Agilent SureScan Microarray Scanner (Agilent Technologies). Исследование выполнено на базе Центра коллективного пользования научно-исследовательским оборудованием и экспериментальным биологическим материалом «Медицинская геномика» НИИ медицинской генетики Томского НИМЦ.

Первичная обработка изображений микрочипов выполнялась в программе Feature Extraction v.11. Для увеличения надежности получаемых результатов использовался принцип конкордантности результатов с применением двух методов: ADM-2, реализованного в программной оболочке Agilent Genomic Workbench v.7 (Agilent Technologies), и метода циркулярной бинарной сегментации (CBS, circular binary segmentation), реализованного в программном пакете DNACopy (Bioconductor) статистической среды R [9,10]. CNV, выявленные у каждого пациента с использованием этих двух алгоритмов, были объединены по границам внешних проб, фланкирующих перестройку. За CNV принимались отклонения от общего централизованного геномного профиля с порогом $\log_2 \pm 0,25$ для ADM-2 и $\log_2 \pm 1,0$ для CBS алгоритмов. CNV имели размер не менее 1 т.п.н. и не менее 3-х последовательных CGH-проб.

Для аннотации выявленных CNV и генов, локализованных в их области, использовались базы данных: Database of Genomic Variants (DGV, сборка NCBI37, hg19), NCBI RefSeq, HuGE Navigator Database, DECIPHER и ENCODE. Анализ наибольшей представленности генов, локализованных в области CNV, проводился с использованием программы Web-based GENE SeT AnaLysis Toolkit [11]. Категории KEGG установлены методом гипергеометрического статистического анализа, где уровни значимости скорректированы с использованием поправки Беньямини—Хохберга.

Результаты и обсуждение

В настоящем исследовании при проведении скрининга CNV методом aCGH в лейкоцитах у пациентов с ИБС идентифицировано 90 вариаций. На одного пациента в среднем приходилось $38,4 \pm 7,5$ CNV. Размер CNV варьировал от 4,4 т.п.н. до 2 млн п.н. (в среднем 217 т.п.н.). Большинство CNV представляло собой уменьшение числа копий участков ДНК — 57 (63%). Пятьдесят пять (61%) из 90 CNV выявлены более чем у одного больного. Среди идентифицированных вариаций числа копий участков ДНК 72 (80%) содержали гены, согласно базе данных NCBI RefSeq.

При функциональной аннотации генов, картированных в областях CNV, выявлено, что их белковые продукты участвуют в иммуновоспалительном ответе (*GBP3*, *IGSF3*, *PRKRA*, *BTNL3*, *АРОВЕС3В*, *АРОВЕС3А*, *LOC391322*, а также семейство генов *DEFB*, $p_{FDR} < 0,05$), обеспечивают функционирование обонятельных рецепторов (*OR2*, *OR4* и *OR11*, $p_{FDR} < 0,05$). Кроме того, белковые продукты генов, подверженных вариациям числа копий участков ДНК, вовлечены в метаболизм различных субстратов (*AMY2B*, *CELA2A*, *CELA2B*, *UGT2B17*, *UGT2B28*, *GSTT1*, $p_{FDR} < 0,05$). Этот результат согласуется с тем, что наиболее часто CNV встречаются в регионах генома, содержащих гены, белковые продукты которых отвечают за взаимодействие с факторами внешней среды [11,12].

Согласно литературным данным, связаны с атеросклерозом и его факторами риска были гены *GBP3*, *GSTT1*, *LOC391322* и *AMY2B*. В лейкоцитах у 7 пациентов с ИБС в локусе 1p22.2 идентифицировано увеличение числа копий размером 4,4 т.п.н., захватывающее ген *GBP3*. Его продукт принадлежит семейству гуанилат-связывающих белков. Их функция заключается в защите против вирусных и бактериальных инфекций. Экспрессия данных генов происходит в ответ на выделение интерферонов. Белки, включая *GBP3*, могут участвовать в активации макрофагов при атеросклерозе. Увеличение экспрессии гена *Gbp3* выявлено в пенных клетках атеросклеротических бляшек синуса аорты на поздних стадиях патологического процесса у мышей *ApoE*^{-/-}, находящихся на «западной диете» в течение 14 недель [13, 14].

Обнаруженное в лейкоцитах у одного пациента с ИБС уменьшение числа копий участков ДНК в хромосомном субсегменте 22q11.23, размером 28,3 т.п.н., затрагивает два гена — *GSTT1* и *LOC391322*. Уменьшение числа копий участков ДНК в гене *GSTT1* приводит к отсутствию активности фермента глутатион-S-трансферазы, ослаблению защиты клетки от влияния окислительного стресса и формированию предрасположенности к артериальной гипертензии и атеросклерозу [15, 16]. В то же время, ген *LOC391322* или *DDTL* (D-dopachrome tautomerase-like) кодирует белок из семейства MIF (ингибирующий фактор миграции макрофагов). Данные белки являются провоспалительными цитокинами, которые обеспечивают продукцию и освобождение различных воспалительных молекул в Т-клетках, моноцитах, в том числе и в клетках атеросклеротических бляшек. Причем увеличение экспрессии гена *MIF* происходит в клетках атеросклеротических бляшек при нарастающей степени тяжести патологического процесса и его нестабильности [17, 18]. Показано, что уровень MIF в сыворотке может быть прогностическим маркером в отношении течения острого ишемического инсульта и положительно коррелирует с размером области поражения головного мозга и неблагоприятным исходом [19]. В связи с этим уменьшение числа копий участков ДНК в двух генах (*GSTT1* и *LOC391322*) может с одной стороны, через снижение защиты клеток от окислительного стресса предрасполагать к заболеванию, а с другой, уменьшая продукцию и освобождение провоспалительных цитокинов, обеспечивать его более благоприятный исход. Согласно базе данных DGV, в локусе 22q11.23 часто встречаются микроструктурные перестройки в виде увеличения и уменьшения числа копий участков ДНК. По-видимому, противоречивые ассоциации между делецией в гене *GSTT1* и риском гипертензии, а также атеросклеротическим поражением коронарных и сонных артерий могут быть связаны с типом и размером хромосомной перестройки в данном локусе.

Также в настоящем исследовании у 7 пациентов в лейкоцитах выявлена делеция в регионе 1p21.1, которая затрагивает ген *AMY2B*, полиморфизм которого ассоциирован с сахарным диабетом 2 типа [20]. Белковый продукт данного гена — альфа-амилаза поджелудочной железы катализирует начальный этап расщепления крахмала и гликогена пищи. Изменение числа копий участков ДНК в гене *AMY2B* коррелирует с уровнем амилазы в сыворотке, причем высокая активность амилазы улучшает гомеостаз глюкозы и снижает риск развития метаболического синдрома, как фактора риска ИБС [21].

Выводы

В результате настоящего исследования впервые в лейкоцитах у пациентов с ИБС картированы регионы генома с CNV с использованием aCGH. Существенная

часть идентифицированных CNV — редкие по частоте и небольшие по размеру. Белковые продукты генов, расположенных в области CNV, участвуют в иммуновоспалительном ответе, обеспечивают функционирование обонятельных рецепторов, а также ферментов метаболизма. Гены, картированные в области CNV в хромосомных субсегментах 1p22.2 (*GBP3*), 1p21.1 (*AMY2B*) и 22q11.23 (*GSTT1*, *LOC391322*) связаны с атеросклерозом и его факторами риска. Однако большая часть идентифицированных CNV не была ранее ассоциирована с атеросклерозом и его факторами риска, в связи с этим она имеет потенциал для раскрытия доли «недостающей наследуемости» ИБС, а также молекулярных мишеней для профилактики, диагностики и лечения данной патологии.

Список литературы

1. Bjorkegren JLM, Kovacic JC, Dudley JT, Schadt EE. Genome-Wide Significant Loci: How Important Are They? *J Am Coll Cardiol*. 2015;65(8):830-845. doi:10.1016/j.jacc.2014.12.033.
2. Eichler EE, Flint J, Gibson G, et al. Missing heritability and strategies for finding the underlying causes of complex disease. *Nat Rev Genet*. 2010;11(6):446-450. doi:10.1038/nrg2809.
3. Pollex RL, Hegele RA. Copy Number Variation in the Human Genome and Its Implications for Cardiovascular Disease. *Circulation*. 2007;115(24):3130-3138. doi:10.1161/CIRCULATIONAHA.106.677591.
4. Manolio TA, Collins FS, Cox NJ, et al. Finding the missing heritability of complex diseases. *Nature*. 2009;461(7265):747-753. doi:10.1038/nature08494.
5. Shia W-C, Ku T-H, Tsao Y-M, et al. Genetic copy number variants in myocardial infarction patients with hyperlipidemia. *BMC Genomics*. 2011;12(Suppl 3):S23. doi:10.1186/1471-2164-12-S3-S23.
6. Costelloe SJ, El-Sayed Moustafa JS, Drenos F, et al. Gene-targeted analysis of copy number variants identifies 3 novel associations with coronary heart disease traits. *Circ Cardiovasc Genet*. 2012;5(5):555-560. doi:10.1161/CIRCGENETICS.111.961037.
7. Kathiresan S., Voight B.F. et al., Myocardial Infarction Genetics Consortium, Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants // *Nat. Genet*. 2009. V. 41. № 3. P. 334-341. doi:10.1038/ng.327.
8. R Development Core Team. R: A Language and Environment for Statistical Computing. R Found Stat Comput Vienna Austria. 2016;0:{ISBN} 3-900051-07-0. doi:10.1038/sj.hdy.6800737.
9. Olshen AB, Venkatraman ES, Lucito R, Wigler M. Circular binary segmentation for the analysis of array-based DNA copy number data. *Biostatistics*. 2004;5(4):557-572. doi:10.1093/biostatistics/kxh008.
10. Wang J, Duncan D, Shi Z, Zhang B. WEB-based GENE SeT Analysis Toolkit (WebGestalt): update 2013. *Nucleic Acids Res*. 2013;41(Web Server issue):W77-83. doi:10.1093/nar/gkt439.
11. Ionita-Laza I, Rogers AJ, Lange C, Raby BA, Lee C. Genetic association analysis of copy-number variation (CNV) in human disease pathogenesis. *Genomics*. 2009;93(1):22-26. doi:10.1016/j.ygeno.2008.08.012.
12. Girirajan S, Campbell CD, Eichler EE. Human Copy Number Variation and Complex Genetic Disease. *Annu Rev Genet*. 2011;45(1):203-226. doi:10.1146/annurev-genet-102209-163544.
13. Han BH. Interferon-gamma and lipopolysaccharide induce mouse guanylate-binding protein 3 (mGBP3) expression in the murine macrophage cell line RAW264.7. *Arch Pharm Res*. 1999;22(2):130-136. http://www.ncbi.nlm.nih.gov/pubmed/10230502.
14. Goo YH, Son SH, Yechoor VK, Paul A. Transcriptional profiling of foam cells reveals induction of guanylate-binding proteins following western diet acceleration of atherosclerosis in the absence of global changes in inflammation. *J Am Heart Assoc*. 2016;5(4). doi:10.1161/JAHA.115.002663.
15. Eslami S, Sahebkar A. Glutathione-S-transferase M1 and T1 null genotypes are associated with hypertension risk: A systematic review and meta-analysis of 12 studies. *Curr Hypertens Rep*. 2014;16(6). doi:10.1007/s11906-014-0432-1.
16. Song Y, Shan Z, Luo C, et al. Glutathione S-Transferase T1 (GSTT1) Null Polymorphism, Smoking, and Their Interaction in Coronary Heart Disease: A Comprehensive Meta-Analysis. *Hear Lung Circ*. 2017;26(4):362-370. doi:10.1016/j.hlc.2016.07.005.
17. Schmeisser A, Marquetant R, Illmer T, et al. The expression of macrophage migration inhibitory factor 1alpha (MIF 1alpha) in human atherosclerotic plaques is induced by different proatherogenic stimuli and associated with plaque instability. *Atherosclerosis*. 2005;178(1):83-94. doi:10.1016/j.atherosclerosis.2004.08.038.
18. Prencipe G, Auriti C, Inglese R, Gallusi G, Dotta A, De Benedetti F. The macrophage migration inhibitory factor -173G/C polymorphism is not significantly associated with necrotizing enterocolitis in preterm infants. *J Pediatr Surg*. 2013;48(7):1499-1502. doi:10.1016/j.jpedsurg.2013.01.004.
19. Li Y-S, Chen W, Liu S, Zhang Y-Y, Li X-H. Serum macrophage migration inhibitory factor levels are associated with infarct volumes and long-term outcomes in patients with acute ischemic stroke. *Int J Neurosci*. 2017;127(6):539-546. doi:10.1080/00207454.2016.1211648.
20. Bailey JNC, Lu L, Chou JW, et al. The Role of Copy Number Variation in African Americans with Type 2 Diabetes-Associated End Stage Renal Disease. *J Mol Genet Med*. 2013;7(2013):61. doi:10.4172/1747-0862.1000061.
21. Nakajima K. Low serum amylase and obesity, diabetes and metabolic syndrome: A novel interpretation. *World J Diabetes*. 2016;7(6):112. doi:10.4239/wjd.v7.i6.112.